Non-Dominated Sorting Genetic Algorithm Based on Reinforcement Learning to Optimization of Broad-Band Reflector Antennas Satellite

Teodoro C. Bora[1], Luiz Lebensztajn[2], Leandro dos S. Coelho[1]
[1] Automation and Systems Laboratory, Industrial and Systems Engineering Graduate Program, Pontifical Catholic University of Paraná, Rua Imaculada Conceição 1155, Zip code: 80215-901, Curitiba, PR, Brazil
[2] Laboratório de Eletromagnetismo Aplicado, LMAG-PEA, Escola Politécnica da Universidade de São Paulo Av. Prof. Luciano Gualberto 158, Zip code: 05508-900, São Paulo, SP, Brazil
E-mails: teodoro.bora@pucpr.br, leb@pea.usp.br, leandro.coelho@pucpr.br

Abstract — This paper aims to provide an improved NSGA-II (Non-Dominated Sorting Genetic Algorithm) which incorporates a parameter-free self-tuning approach by reinforcement learning technique, called Non-Dominated Sorting Genetic Algorithm Based on Reinforcement Learning (NSGA-RL). The proposed method is particularly compared with the classical NSGA-II when applied to a satellite coverage problem. Furthermore, the optimization results are compared with other results in the with others multiobjective optimization methods.

I. INTRODUCTION

The study of satellite broad-cast communication had been done in several works, [1],[3],[4]. Basically, the main objective of this class of problem is to reach a maximum gain and illumination uniformity inside a prescribed region [1],[4].

In this paper, for the optimization task of satellite broad-cast communication, two multiobjective optimization techniques based on genetic algorithms are validated. The techniques are: (i) classical Non-Dominated Sorting Genetic Algorithm (NSGA-II), proposed in [3], and (ii) a new improved NSGA-II called Learner Non-Dominated Sorting Genetic Algorithm Based on Reinforcement Learning (NSGA-RL).

The classical NSGA-II uses an elitist selection through its domination sort algorithm and, also uses a parameter-free diversity metric denominated crowding distance. On the other hand, the proposed NSGA-RL is a parameter-free bio-inspired algorithm based on the classical one proposed in [2]. However, it is based on reinforcement learning techniques to self-tuning its probabilities and indices.

The rest of this paper is organized as follows: In Section II, the basis of satellite coverage problem is presented. In sections III and IV, the optimization procedure of the classical NSGA-II and NSGA-RL are detailed, respectively. Finally, the achieved numerical results by NSGA-II and NSGA-RL are compared in Section V with others results in the recent literature. To finish, Section VI discusses the results and possible advances.

II. THE OPTIMIZATION PROBLEM DEFINITION

The complete mathematical model of the satellite coverage problem is well described and discussed in [1],[3],[4], and it will be omitted here. Shortly, it could be stated that an antenna in geosynchronous orbit must illuminate a target on the planet surface as follows: the maximum gain and illumination uniformity inside a given region should be, as close as possible, reached [4]. These two objectives are evaluated using only one objective function, which is considered in three different frequencies, [1], instead of one as in [3], and subject to a minimum at some sub region [4]. This formulation is stated in equation (1), as proposed in [4], where some control points p are spread over the target region P, $Q \in P$ and G is the gain radiation (dBi).

$$\begin{aligned} &minimize \quad f(x) = -\operatorname{mean}_{p_i \in P} G(x, p_i) \qquad (1)\\ &subject\ to \quad g_j(x) = G_{min} - G(x, q_j) \leq 0, \ \ \forall q_j \in Q \end{aligned}$$

The problem has 38 design variables, which came from the 25 control points in Bernstein-Bézier surfaces, which defines the reflector shape. The other variables are linked to the feed position and rotation parameterization [4].

III. CLASSICAL NSGA-II

The NSGA-II was built as its classical structure as proposed in [2]. The crossover operator method is the simulated binary crossover (SBX), as shown in [6]. On the other hand, the mutation is performed by the polynomial mutation method stated also in [6]. These methods require 2 parameters called distribution indexes, which define the spread of the created solutions. In this work, these indices were set as 40. Besides, crossover and mutation occurrence probability are 0.9 and 0.03 (per gene), respectively. The tournament section method is obtained throughout the matting pool technique, and the function evaluation by the same Matlab scripts used in [4] and available in [5].

IV. THE REINFORCEMENT LEARNING SELF-ADAPTATION APPROACH FOR THE NSGA-II

It is widely known that the quality of the Pareto frontier reached by a meta-heuristic, like NSGA-II, is highly dependent on its parameter set. This statement is proved by the No Free-Lunch Theorem [7]. Therefore, the self-adaptive features are studied in the modern papers of nature-inspired optimization algorithm field. The NSGA-RL is one of them, based on a greedy classic reinforcement learning technique.

The main idea in NSGA-RL is to incorporate learning features and, consequently, knowledge to the population. Thus, based on the past generations, it is possible to make some decisions concerning the variation of the crossover and mutation. The knowledge is stored based on the domination rank elitism structure. For each new generated offspring, the population is feedback with a reward, which is straight related to the success of this crossover and mutation operation.

In this way, the population learns, by try and error procedure, which is the best values for parameters, whether exist ones, to adapt themselves for non-stationary problems. These parameters are the crossover and mutation probabilities and indexes as defined in Section III. For each parameter, a tridimensional matrix is then built. The matrix indexes are defined by the parents' rank and a set of discrete values of the parameter. The matrix is then filled with the sum of the past rewards after the parameter application.

Then, for a given couple of parents, their ranks and related rewards, allows us to define the best value for the concerning parameter. The NSGA-RL uses a simple greedy selection reinforcement learning technique [8] with maximum/minimum saturation structure. Then, when one parameter has more rewards than any other it is selected, otherwise, when a draw occurs, the selection is done randomly among the greatest ones.

V. OPTIMIZATION RESULTS

The objective function and its constraints (the gain should be greater than 35dBi in each control point in a given region), shown by (1), were optimized by NSGA-II and NSGA-RL. Table I shows some results for both methods and for the multiobjective Immune Systems (IS) and Particle Swarm Optimization (PSO) algorithms for 10 simulations. The IS and PSO results were extracted from [4] and every metric was performed with 10 runs. Table I gives a straight comparison tool to determine the effectiveness of the NSGA-II and NSGA-RL approaches.

TABLE I
COMPARISON OF RESULTS FOR SEVERAL APPROACHES WITH 10 RUNS. IS AND PSO EXTRACTED FROM [4]

Approach	Brazil		China		USA	
	Mean	Best	Mean	Best	Mean	Best
NSGA-II Penalty	29.88 ± 0.53	30.35	31.53 ± 0.27	31.78	31.80 ± 0.25	32.26
NSGA-RL Penalty	30.06 ± 0.29	30.34	31.42 ± 0.35	31.76	31.84 ± 0.32	32.30
IS Penalty	29.9 ± 0.3	30.34	32.0 ± 0.2	32.17	32.8 ± 0.1	32.95
PSO Penalty	28.8 ± 0.7	29.54	31.2 ± 0.8	31.93	31.9 ± 0.8	32.96

In Table II the classic NSGA-II is compared with the NSGA-RL approach with 30 runs to guarantee a statistical relevance. Besides, for both NSGA-II and NSGA-RL, the number of individuals was set to 100. NSGA-II ran for 300 generations, while for NSGA-RL only for 200 had been chosen. Unfortunately, in [4], it was not found in how well the minimum of 35dBi were satisfied. It is so important to comment that obeying these constraints implies in a straight minimization in the mean gain. For both the NSGA-II and NSGA-RL these constraints were well satisfied, what had a great effect in the best solution mean gain. In all cases, the constraints were dealt as penalty functions.

Besides, the 30 run comparison between the NSGA-II and the NSGA-RL showed that this self-adaptive approach has a good performance, when compared with the NSGA-II. In this problem, we realize that the constraints play a very important role. We have observed that the gain are always greater than, or at least equal to 35dBi in each control point in a given region, for all countries. The worst case is a control point in Brazil: 35.002 dBi.

TABLE II
COMPARISON OF RESULTS AFTER 30 RUNS

Approach	Brazil		China		USA	
	Mean	Best	Mean	Best	Mean	Best
NSGA-II Penalty	29.90 ± 0.47	30.43	31.42 ± 0.37	31.81	31.79 ± 0.22	32.26
NSGA-RL Penalty	30.0 ± 0.28	30.38	31.47 ± 0.26	31.78	31.87 ± 0.38	32.32

When we analyze the computational time, the NSGA-RL, due to its adaptive nature, is, as expected, slower than the NSGA-II. From our experience, the NSGA-RL time computation could be twice the NSGA-II time, for some analytical benchmarks. Nevertheless, when solving the optimization of antennas satellite, the function evaluation takes around 99.6% of the total time cost for both methods, so the NSGA-RL adaptive process is well suited to this class of problem.

VI. CONCLUSION

As could be seen the parameter-free variation of the classical NSGA-II performed well in relations to others methods. Besides, the NSGA-RL was better in many aspects whether compared with NSGA-II and with the advantage of non-parameter setting and a little number of generations run. As consequence fewer evaluations of objective functions and improved speed are obtained using the NSGA-RL.

It is important to note that the NSGA-RL has a small standard deviation in the best solution, which can be used to say that it appears to be more robust than other methods, in other words, there in not a so big difference in the best solution between the runs.

REFERENCES

[1] A. C. Lisboa, D. A. G. Vieira, J. A. Vasconcelos, R. R. Saldanha, R. H. C. Takahashi, "Multiobjective Shape Optimization of Broad-Band Reflector Antennas Using the Cone of Efficient Directions Algorithm," *IEEE Transactions on Magnetics*, vol. 42, no. 4, pp. 1223-1226, 2006.

[2] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, "A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, 2002.

[3] A. C. Lisboa, D. A. G. Vieira, J. A. Vasconcelos, R. R. Saldanha, R. H. C. Takahashi, "Decreasing Interference in Satellite Broadband Communication Systems Using Modeled Reflector Antennas," *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 958-961, 2008.

[4] D. A. G. Vieira, A. C. Lisboa, R. R. Saldanha, R. H. C. Takahashi, "An Enhanced Ellipsoid Method for Electromagnetic Devices Optimization and Design", *IEEE Transactions on Magnetics*, vol. 46, no. 8, pp. 2843-2851, 2010.

[5] Enacom website, *http://www.enacom.com.br/reflector.html*

[6] K. Deb, "*Multi-Objective Optimization using Evolutionary Algorithms*", Wiley-Interscience Series in Systems and Optimization, 2001.

[7] D. H. Wolpert, W. G. Macready, "No Free Lunch Theorems for Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67-82, 1997.

[8] R. S. Sutton, A. G. Barto, "*Reinforcement Learning: An Introduction*", MIT Press, MA, 1998.